



Transformation d'un dispositif multimédia webcam-écran en un scanner 3D

Yvain Quéau, Richard Modrzejewski, Pierre Gurdjos, Jean-Denis Durou

► To cite this version:

Yvain Quéau, Richard Modrzejewski, Pierre Gurdjos, Jean-Denis Durou. Transformation d'un dispositif multimédia webcam-écran en un scanner 3D. COMpression et REprésentation des Signaux Audiovisuels (CORESA), Nov 2014, Reims, France. pp.0. hal-01120851

HAL Id: hal-01120851

<https://hal.science/hal-01120851>

Submitted on 26 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 13272

To cite this version :

Quéau, Yvain and Modrzejewski, Richard and Gurdjos, Pierre and Durou, Jean-Denis *Transformation d'un dispositif multimédia webcam-écran en un scanner 3D*. (2014) In: COMpression et REprésentation des Signaux Audiovisuels (CORESA), 26 November 2014 - 27 November 2014 (Reims, France). (Unpublished)

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Transformation d'un dispositif multimédia webcam-écran en un scanner 3D

Y. Quéau R. Modrzejewski P. Gurdjos J.-D. Durou

IRIT, UMR CNRS 5505, Toulouse

Résumé

Nous étudions un dispositif de scannage 3D constitué d'un couple webcam-écran, où l'écran est utilisé comme source lumineuse. Ceci permet de transformer en scanner 3D n'importe quel dispositif multimédia comprenant ces deux éléments (ordinateurs portables, smartphones, tablettes etc.). Un protocole d'étalonnage simplifié est introduit, pour lequel nous démontrons que deux prises de vue sont suffisantes. Une fois cet étalonnage géométrique effectué, nous montrons que le dispositif étudié permet d'effectuer la reconstruction 3D sans ambiguïté, grâce à la technique de stéréophotométrie.

We study a multimedia system composed of both a webcam and a screen, where the screen is considered as a light source. Such a system allows one to turn a laptop, a smartphone or a tablet into a 3D-scanner. We propose a simplified calibration procedure which requires as few as two input images, and demonstrate that such a system allows one to get a 3D-reconstruction without ambiguity, using the photometric stereo technique.

Mots clé : étalonnage, webcam, écran LCD, reconstruction 3D, stéréophotométrie.

1. Introduction

L'utilisation de dispositifs multimédia comprenant à la fois un écran et une webcam est devenue incontournable. Dans de nombreux terminaux numériques tels que les ordinateurs portables, les smartphones ou les tablettes, ces deux entités coexistent. Et lorsque cela n'est pas le cas, il est toujours possible d'attacher la webcam à l'écran par une pince (cf. figure 1).

En détournant l'écran de son usage standard pour le transformer en source lumineuse [FY07], on dispose d'un système de reconstruction 3D par stéréophotométrie : m images de la scène sont acquises sous le même point de vue (le système est supposé immobile, tout comme la scène à reconstruire), mais sous m éclairages différents, contrôlés par l'affichage de motifs connus sur l'écran du dispositif. Pour pouvoir caractériser géométriquement l'éclairage incident relativement à la scène, il est toutefois nécessaire d'avoir étalonné géométriquement le système, c'est-à-dire de connaître la pose de la webcam relativement à l'écran.

Notre contribution est double : nous introduisons d'abord une procédure d'étalonnage géométrique simplifiée, dans le cas particulier où le mouvement de la webcam est limité à une rotation autour d'un axe parallèle à l'axe horizontal de l'écran (cf. figure 1); nous montrons ensuite que, grâce à

cet étalonnage, le dispositif multimédia étudié permet de garantir l'unicité de la solution au problème de la reconstruction 3D par stéréophotométrie, là où la grande majorité des méthodes existantes ne fournissent de solution qu'à une ambiguïté près.

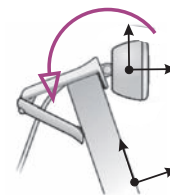


Figure 1: Le dispositif consiste en un écran auquel est attachée une caméra (webcam). La caméra possède quatre degrés de liberté relativement à l'écran, qui correspondent aux trois coordonnées de son centre et à une rotation autour de l'axe horizontal de l'écran. (source : www.logitech.com/en-us/support/hd-webcam-c270).

2. Étalonnage géométrique du système

Dans [RBN10], il est établi que l'image d'un objet obtenue par réflexion sur un miroir plan est géométriquement équivalente à l'image de cet objet obtenue par une caméra virtuelle située derrière le miroir, qui aurait les mêmes paramètres internes que la caméra réelle et dont la pose serait symétrique, par rapport au miroir, à celle de la caméra réelle

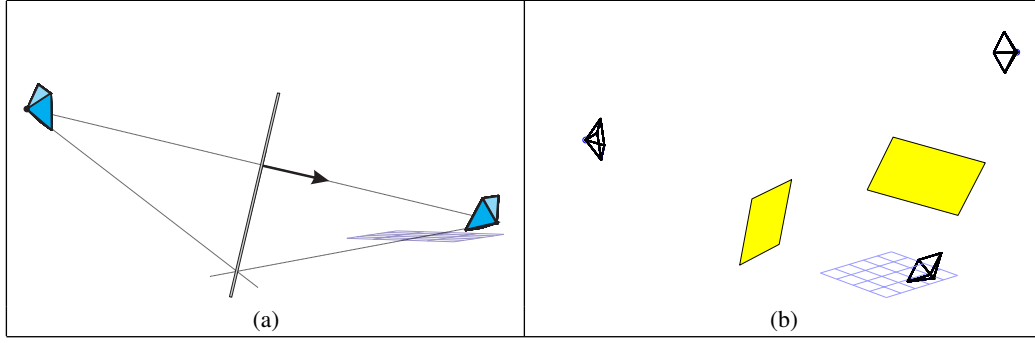


Figure 2: (a) En plaçant un miroir face à la caméra réelle (à droite), on définit une caméra virtuelle (à gauche), symétrique de la caméra réelle par rapport au plan du miroir. (b) En utilisant plusieurs poses du miroir (en jaune), on définit autant de caméra virtuelles que de poses.

(cf. figure 2-a). Nous allons utiliser cette propriété pour l'étalonnage géométrique du dispositif, et montrer qu'en utilisant le miroir pour réfléchir un motif affiché à l'écran, l'étalonnage géométrique du système (paramètres internes de la caméra et pose relativement à l'écran) peut être effectué analytiquement, à partir d'un minimum de deux prises de vue.

2.1. Notations

Dans ce travail, nous considérons que le repère tridimensionnel de référence est un repère orthonormé attaché à l'écran, orienté comme l'indique la figure 1 : l'axe des y est parallèle aux colonnes de l'écran, tandis que l'axe des z est orthogonal à l'écran.

Nous modélisons la caméra par un « trou d'épingle » de type CCD [HZ03]p.156, caractérisé par la position $W = (X_w, Y_w, Z_w)$ de son centre optique et l'orientation de l'axe optique. Motivés par l'observation des terminaux tels que les tablettes tactiles ou les webcams que l'on peut fixer sur un écran, nous nous limitons au cas suivant :

Hypothèse 1 La caméra est fixée à l'écran par une pince, qui limite son mouvement à une rotation autour d'un axe parallèle à l'axe horizontal de l'écran (cf. figure 1).

La caméra a donc quatre degrés de liberté qui correspondent à l'angle θ de la rotation autour de cet axe et aux trois coordonnées cartésiennes X_w, Y_w, Z_w de son centre optique.

D'après le modèle de caméra CCD, la webcam est décrite par sa matrice de projection, de la forme :

$$P = KR [1 \mid -\mathbf{t}_w] \quad (1)$$

où K est la matrice des paramètres internes de la webcam, I est la matrice unité et :

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \quad \text{et} \quad \mathbf{t}_w = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \quad (2)$$

Le miroir utilisé est supposé parfaitement plan. Ses n poses définissent n miroirs virtuels \mathcal{M}_i , représentés géométriquement par leurs plans de support. Le plan de support du miroir virtuel \mathcal{M}_i est situé à la distance d_i de l'origine du repère considéré, et son vecteur normal est noté \mathbf{n}_i

($\|\mathbf{n}_i\| = 1$). Il est donc caractérisé par le vecteur de coordonnées $[\mathbf{n}_i^\top, d_i]^\top$.

Enfin, nous définissons une caméra virtuelle par réflexion de la webcam par rapport à chaque pose du miroir, qui est décrite par une matrice de projection de la forme :

$$P_i = K_i R_i [1 \mid -\mathbf{t}_w] \quad (3)$$

Le système étudié a les propriétés suivantes, $\forall i \in [1, n]$:

$$K = K_i \quad (4)$$

$$P = P_i S_i \quad (5)$$

où :

$$S_i = \begin{bmatrix} 1 - 2\mathbf{n}_i^\top & -2d\mathbf{n}_i \\ \mathbf{0}_3^\top & 1 \end{bmatrix} \quad (6)$$

est la matrice de la réflexion par rapport au miroir \mathcal{M}_i .

Le problème d'étalonnage que nous devons résoudre peut alors être formalisé de la façon suivante :

Problème 2 À partir des poses des caméras virtuelles, sous l'hypothèse 1, déterminer la pose de la webcam dans un repère orthonormé attaché à l'écran.

Nous allons montrer que le problème 2 admet une solution analytique unique, à partir d'un minimum de deux poses du miroir. En particulier, nous allons établir que la composante rotationnelle de la pose de la webcam est solution d'un système linéaire homogène, et que la composante translationnelle peut être obtenue en résolvant un problème d'intersection de droites.

2.2. Estimation des paramètres intrinsèques

Commençons par l'estimation de la matrice K des paramètres intrinsèques. Ces paramètres peuvent être estimés à partir de $m \geq 2$ images d'une mire d'étalonnage (par exemple un damier) posée sur un support plan, prises sous différentes orientations.

Le miroir peut être utilisé pour simuler ce protocole : une mire d'étalonnage affichée à l'écran est réfléchi sur le

miroir. En changeant l'orientation du miroir, tout se passe comme si l'on utilisait des mires virtuelles directement visibles depuis la caméra. Les méthodes classiques d'étalonnage peuvent alors être utilisées. Dans notre implémentation, nous utilisons la méthode de Bouguet [Bou04].

Les caméras virtuelles ayant les mêmes paramètres internes que la webcam, cf. (4), ceci permet également d'estimer les matrices K_i , ainsi que les poses des caméras virtuelles relativement au miroir (matrices R_i et vecteurs \mathbf{t}_i).

2.3. Estimation des paramètres extrinsèques

Intéressons-nous maintenant au calcul de la pose de la webcam, représentée par le produit $R [1 \mid -\mathbf{t}_w]$.

Composante rotationnelle de la pose de la webcam. Il s'agit d'estimer la matrice R . Par souci de clarté, nous omettons momentanément les indices i pour les caméras virtuelles, car il est possible d'obtenir une solution unique en R pour chacune d'elles : l'estimation peut ensuite être rendue robuste en moyennant les résultats obtenus pour les différentes caméras virtuelles.

Considérons la sous-matrice obtenue en supprimant les troisièmes ligne et colonne de S :

$$\tilde{S} = I - 2\mathbf{n}\mathbf{n}^\top \quad (7)$$

On montre facilement que :

$$\det \tilde{S} = -1 \quad (8)$$

De l'égalité (5), on déduit que :

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \sim R\tilde{S} \quad (9)$$

où \sim désigne l'égalité projective.

En exploitant l'égalité (9), on utilise pour estimer \tilde{S} les six contraintes linéaires suivantes par rapport aux six éléments de la matrice (symétrique) \tilde{S} :

$$\underbrace{\begin{bmatrix} 0 & -R_{21} & R_{31} & -R_{22} & R_{32} - R_{23} & R_{33} \\ 0 & R_{31} & R_{21} & R_{32} & R_{22} + R_{33} & R_{23} \\ 0 & R_{11} & 0 & R_{12} & R_{13} & 0 \\ 0 & 0 & R_{11} & 0 & R_{12} & R_{13} \\ R_{21} & R_{21} & R_{23} & 0 & 0 & 0 \\ R_{31} & R_{32} & R_{33} & 0 & 0 & 0 \end{bmatrix}}_C \underbrace{\begin{pmatrix} S_{11} \\ S_{12} \\ S_{13} \\ S_{22} \\ S_{23} \\ S_{33} \end{pmatrix}}_X = \mathbf{0}_6 \quad (10)$$

D'un point de vue théorique, le système obtenu est compatible et admet une solution exacte car la matrice C est, en général, une matrice d'ordre 6 et de rang 5. En présence de bruit sur les données, on a $\text{rang}(C) > 5$, donc une solution au sens des moindres carrés doit être recherchée. Pour éviter la solution triviale $\mathbf{X} = \mathbf{0}_6$, on impose la contrainte $\|\mathbf{X}\| = 1$. Le problème s'écrit alors $\arg \min_{\mathbf{X}} \|\mathbf{C}\mathbf{X}\|^2$ s.c. $\|\mathbf{X}\| = 1$, qui peut être résolu par décomposition en valeurs singulières. Soit \hat{S} la solution pour \tilde{S} , normalisée de telle façon que (8) soit satisfaite. On obtient :

$$\cos \theta = (R\hat{S})_{22} \quad ; \quad \sin \theta = (R\hat{S})_{32} \quad (11)$$

en notant que $R\hat{S}$ est nécessairement de la forme du membre gauche de (9).

Composante translationnelle de la pose de la webcam. Il s'agit d'estimer la translation \mathbf{t}_w du centre optique de la webcam. Ce dernier est situé à l'intersection des droites passant par les centres optiques des n caméras virtuelles, représentés par les vecteurs \mathbf{t}_i , et dirigées par les vecteurs normaux aux plans des miroirs correspondants, représentés par \mathbf{n}_i .

On connaît les vecteurs \mathbf{t}_i et, en utilisant l'égalité (7), on estime les vecteurs \mathbf{n}_i à partir de la décomposition en valeurs singulières de $\frac{1}{2}(I - \hat{S}_i)$:

$$\hat{\mathbf{n}}_i = U(1, 0, 0)^\top \quad \text{si} \quad U\Sigma V^\top = \frac{1}{2}(I - \hat{S}_i) \quad (12)$$

où U et V sont deux matrices orthogonales d'ordre 3 et Σ est la matrice diagonale des valeurs singulières. Il subsiste une ambiguïté sur le signe de $\hat{\mathbf{n}}_i$, qui peut être levée aisément puisqu'on sait que le miroir est orienté face à la caméra.

Chaque droite passant par le centre, de vecteur \mathbf{t}_i , d'une caméra virtuelle numéro i , et dirigée par le vecteur normal \mathbf{n}_i , peut être représentée par une matrice de Plücker [HZ03]p.70. En introduisant les vecteurs de coordonnées homogènes $\tilde{\mathbf{n}}_i = (\mathbf{n}_i^\top, 0)^\top$ et $\tilde{\mathbf{t}}_i = (\mathbf{t}_i^\top, 1)^\top$, cette matrice peut s'écrire :

$$L_i = \tilde{\mathbf{t}}_i \tilde{\mathbf{n}}_i^\top - \tilde{\mathbf{n}}_i \tilde{\mathbf{t}}_i^\top \quad (13)$$

L_i est une matrice d'ordre 4, définie à un facteur près, antisymétrique avec une diagonale nulle, et de rang 2. Elle ne comporte donc que quatre degrés de liberté, comme toute droite de l'espace affine tridimensionnel.

On utilisera comme équations de base du problème d'intersection les équations linéaires fournies par le système homogène suivant :

$$\underbrace{\begin{bmatrix} L_1^* \\ \vdots \\ L_n^* \end{bmatrix}}_D \underbrace{\begin{pmatrix} \mathbf{t}_w \\ 1 \end{pmatrix}}_{\tilde{\mathbf{Y}}} = \mathbf{0}_{4n \times 4} \quad (14)$$

où L_i^* désigne la seconde matrice de Plücker duale à L_i . Il s'ensuit qu'un minimum de deux droites est nécessaire pour estimer la position du centre de la webcam. Pour éviter la solution triviale $\tilde{\mathbf{Y}} = \mathbf{0}_4$, on impose la contrainte $\|\tilde{\mathbf{Y}}\| = 1$. Le problème s'écrit alors $\arg \min_{\tilde{\mathbf{Y}}} \|\mathbf{D}\tilde{\mathbf{Y}}\|^2$ s.c. $\|\tilde{\mathbf{Y}}\| = 1$, qui peut être résolu par décomposition en valeurs singulières.

2.4. Considérations expérimentales

À partir de n images du miroir, on peut retrouver la pose de la caméra réelle par rapport à l'écran. Nous résumons les différentes étapes de l'étalonnage géométrique dans l'algorithme suivant :

Données : n images d'un miroir réfléchissant un écran sur lequel est affichée une mire d'étalonnage.

Résultats :

- (R, \mathbf{t}_w) : pose de la caméra réelle.
- $[\mathbf{n}_i^T, d_i]^\top$, $i \in [1, n]$: vecteurs des n plans de support des miroirs virtuels.

1. Initialisation : estimer les paramètres internes de la caméra réelle et les poses (R_i, \mathbf{t}_i) des n caméras virtuelles [Bou04].

2. Composante rotationnelle :
Pour $k = 1 \dots n$, estimer θ_i depuis R_i , à partir du système linéaire (10) et de l'identification (11).
3. $\theta = \text{médiane}\{\theta_i\}$
4. $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$
5. Composante translationnelle :
Pour $k = 1 \dots n$, calculer la matrice de Plücker L_k par l'équation (13), et sa seconde matrice duale L_k^* .
6. Résoudre le système (14) au sens des moindres carrés et en déduire t_w .

En théorie, deux poses du miroir suffisent à étalonner géométriquement le système. En effet, il s'agit du nombre minimal d'images permettant de réaliser l'étalonnage des paramètres intrinsèques et des poses des caméras virtuelles [Bou04]. Ensuite, l'estimation de la composante rotationnelle peut être théoriquement menée à partir d'une seule pose, mais en pratique il est nécessaire d'en utiliser davantage, afin de garantir une certaine robustesse aux effets de bruit ou de flou de bougé : pour cela, nous effectuons une estimation par pose, avant de calculer la médiane des estimations. Ceci est également valable pour l'étalonnage de la composante translationnelle, qui est théoriquement possible à partir de deux poses du miroir, mais l'estimation sera bien sûr plus robuste si davantage de poses sont utilisées.

En pratique, nous utilisons un minimum de six poses du miroir. Différentes images du miroir, acquises sous différentes poses, sont représentées sur la figure 3, avec une vue de la reconstruction 3D de la scène d'étalonnage.

3. Application : reconstruction 3D par stéréophotométrie

La stéréophotométrie [Woo80] est une technique de reconstruction 3D monoculaire, où la scène est photographiée sous différents éclairages afin d'en extraire le relief et la réflectance. Nous proposons d'utiliser le dispositif multimédia webcam-écran étudié dans le paragraphe précédent comme scanner 3D, à la manière de [FY07, Sch08] : l'écran permet d'afficher successivement m images qui éclairent la scène à reconstruire (cf. figure 4). La webcam capture ces images, desquelles sont extraites hors ligne les caractéristiques géométriques et photométriques de la scène. Ceci est possible grâce au calibrage géométrique du système, qui permet de spécifier géométriquement le flux lumineux incident.

3.1. Modélisation de l'écran vu comme une source lumineuse

Lorsque l'écran est utilisé en mode projection, chacun de ses pixels peut être considéré comme une source lumineuse ponctuelle anisotrope ayant pour direction principale la normale à l'écran [FY07]. Ce dispositif est représenté sur la figure 5.

Soit \mathcal{E} l'ensemble des pixels d'un motif affiché à l'écran, et $\mathbf{X}_\mathcal{E} \in \mathcal{E}$. On suppose que le motif affiché à l'écran est

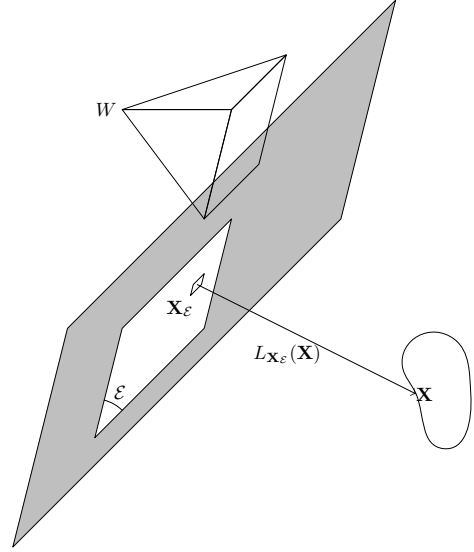


Figure 5: L'écran utilisé comme source lumineuse. Chaque pixel $\mathbf{X}_\mathcal{E}$ de la partie allumée \mathcal{E} de l'écran émet de la lumière dans toutes les directions. La contribution de ce pixel à l'éclairage reçu par un point \mathbf{X} de la scène est notée $L_{\mathbf{X}_\mathcal{E}}(\mathbf{X})$.

uniforme et que tous les pixels réagissent de la même façon, c'est-à-dire que chaque pixel de \mathcal{E} a la même intensité. En tout point $\mathbf{X} = [x_X, y_X, z_X]^T$ de l'espace tel que $z > 0$, la contribution du pixel $\mathbf{X}_\mathcal{E}$ à l'éclairage incident en \mathbf{X} peut être modélisée par la formule suivante :

$$L_{\mathbf{X}_\mathcal{E}}(\mathbf{X}) = \frac{f(\theta)}{\|\mathbf{X} - \mathbf{X}_\mathcal{E}\|^3} (\mathbf{X} - \mathbf{X}_\mathcal{E}) \quad (15)$$

où f représente l'anisotropie de la source : f est en général une fonction décroissante de l'angle θ entre la normale à l'écran et la direction $\mathbf{X} - \mathbf{X}_\mathcal{E}$. Cette fonction peut être calibrée par une procédure spécifique [PSM*14], ou choisie empiriquement. Il a été montré dans [Sch08] que les écrans LCD pouvaient être modélisés par une décroissance en cosinus. Grâce au choix de l'écran comme référence du système de coordonnées, cela s'exprime très simplement par $f(\theta) = \frac{z_X}{\|\mathbf{X} - \mathbf{X}_\mathcal{E}\|}$, ce qui fournit le modèle suivant :

$$L_{\mathbf{X}_\mathcal{E}}(\mathbf{X}) = \frac{z_X}{\|\mathbf{X} - \mathbf{X}_\mathcal{E}\|^4} (\mathbf{X} - \mathbf{X}_\mathcal{E}) \quad (16)$$

3.2. Modélisation photométrique de l'image

Soit \mathcal{S} une surface à reconstruire, placée face au dispositif multimédia webcam-écran. Soit $\mathbf{X} \in \mathcal{S}$ et $\mathbf{N}_\mathbf{X}$ la normale unitaire sortante à la surface au point \mathbf{X} .

Comme les différents pixels émettent des ondes lumineuses incohérentes, le faisceau lumineux incident en \mathbf{X} est égal à la somme des contributions des pixels $\mathbf{X}_\mathcal{E}$. D'après le modèle lambertien, la luminance obtenue dans les canaux rouge, vert et bleu, lorsque la scène est éclairée par le motif

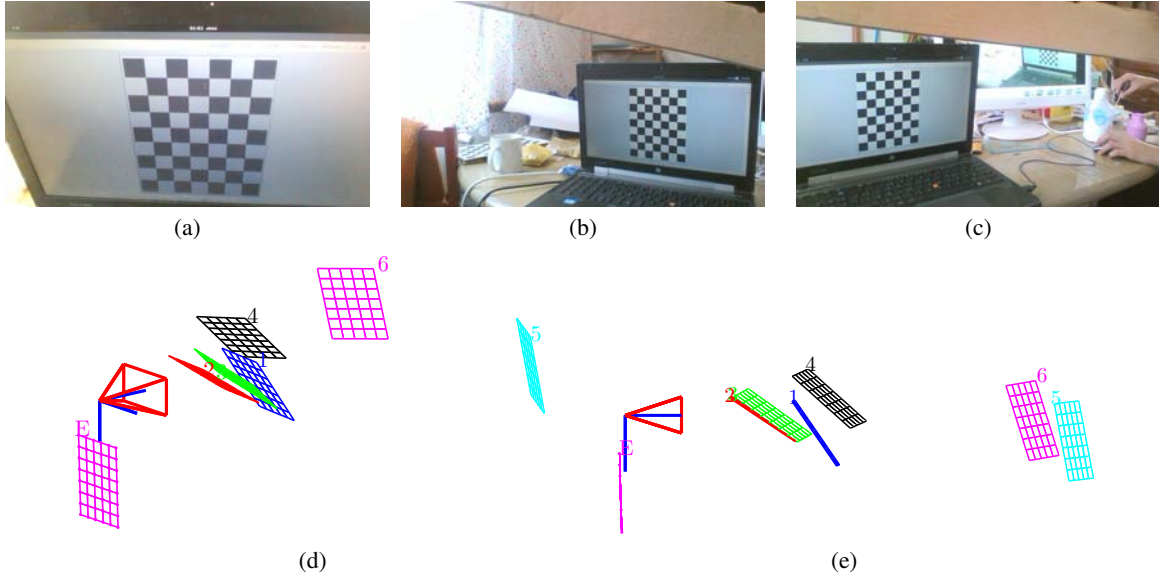


Figure 3: Résultats de la procédure d'étalonnage appliquée à un ordinateur portable avec caméra intégrée. (a-b-c) Trois des images d'étalonnage acquises par la webcam. La mire est directement affichée à l'écran, et réfléchiée par le miroir vers la webcam (le miroir est partiellement visible sur l'image de droite). (d-e) « Scène » reconstruite : les poses des différents miroirs et de la webcam est estimées relativement à l'écran, ce qui permet de spécifier géométriquement l'ensemble de la scène d'étalonnage (les images (a-b-c) correspondent aux poses 1, 5 et 6 de la reconstruction).

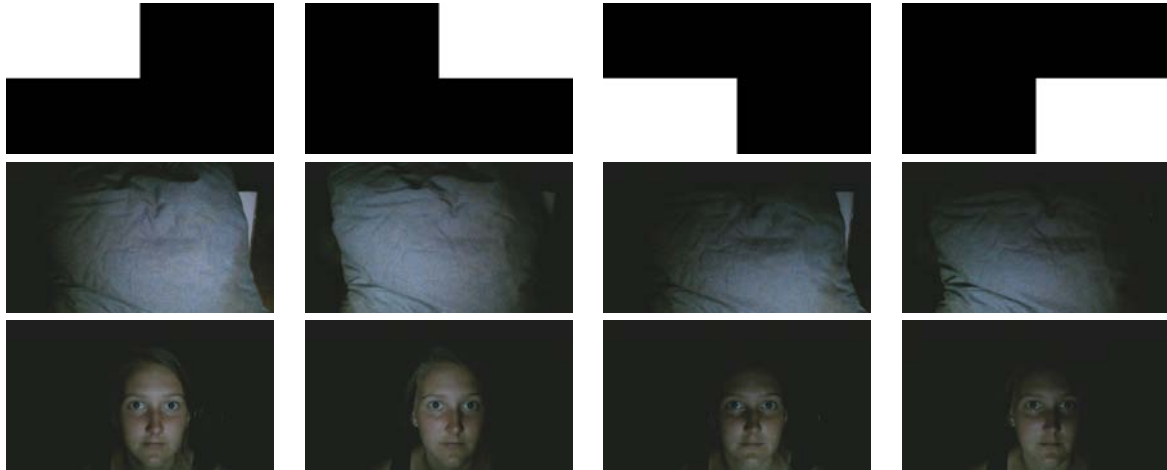


Figure 4: Le dispositif multimédia utilisé comme un scanner 3D : m différents motifs sont affichés successivement à l'écran (première ligne), ce qui permet d'obtenir m images sous différents éclairages (deuxième et troisième lignes). À partir de ces images et de la pose de la caméra relativement à l'écran, on peut estimer un modèle 3D de la surface.

\mathcal{E}_i , vaut :

$$I_{\mathcal{E}}^i(\mathbf{P}\mathbf{X}) = -\rho_{\mathbf{X}}\mathbf{N}_{\mathbf{X}}^{\top} \sum_{\mathbf{X}_{\mathcal{E}} \in \mathcal{E}_i} \mathbf{L}_{\mathbf{X}_{\mathcal{E}}}(\mathbf{X}) \quad (17)$$

$$= -z_{\mathbf{X}} \rho_{\mathbf{X}} \mathbf{N}_{\mathbf{X}}^{\top} \sum_{\mathbf{X}_{\mathcal{E}} \in \mathcal{E}_i} \frac{\mathbf{X} - \mathbf{X}_{\mathcal{E}}}{\|\mathbf{X} - \mathbf{X}_{\mathcal{E}}\|^4} \quad (18)$$

où l'albédo $\rho_{\mathbf{X}}$ représente la proportion de lumière réémise par la surface, relativement au canal de couleur considéré. Le problème de la reconstruction 3D par stéréophotométrie à partir du dispositif webcam-écran peut alors être formalisé de la façon suivante :

Problème 3 À partir de m images I^i , $i = 1 \dots m$, acquises sous m motifs d'éclairage connus \mathcal{E}_i , $i = 1 \dots m$, et de la pose \mathbf{P} de la webcam, estimer les points 3D \mathbf{X} de la surface dans le repère de l'écran, ainsi que l'albédo associé $\rho_{\mathbf{X}}$.

3.3. Inversion du modèle

Afin de résoudre le problème 3, il faut inverser le système formé par les m équations (18) correspondant aux m prises de vue. Cette étape sera détaillée dans un travail futur dédié au problème de la reconstruction 3D par stéréophotométrie en présence de sources lumineuses « étendues », et ne sera



Figure 6: Résultat de la reconstruction 3D à partir du dispositif multimédia, appliquée aux deux séries de 4 images couleur de la figure 4. Le dispositif permet de retrouver à la fois les coordonnées 3D des points de la surface photographiée, et l'albédo relativement à chacun des canaux de couleur. Le dispositif multimédia utilisé est un ordinateur personnel HP EliteBook 8570w avec caméra HD 1280×720 .

donc pas détaillée dans cet article. Notons simplement que, grâce à la présence du terme d'anisotropie, les points X apparaissent explicitement dans le modèle photométrique de l'image, ce qui n'est généralement pas le cas en stéréophotométrie. En conséquence, l'ambiguïté d'échelle liée à l'intégration perspective du champ de normales est évitée, et il est possible d'obtenir la reconstruction 3D sans ambiguïté. Deux exemples de reconstruction 3D, obtenues à partir des jeux d'images de la figure 4, sont présentés sur la figure 6.

4. Conclusion

Dans cet article, nous avons montré qu'il était possible de transformer n'importe quel dispositif multimédia comprenant un écran et une webcam en un scanner 3D capable d'estimer sans ambiguïté le relief de la scène ainsi que l'albédo relativement à chaque canal de couleur.

Pour ce faire, nous avons d'abord introduit une procédure d'étalonnage géométrique spécifique permettant de retrouver la pose de la webcam, relativement à l'écran qui est utilisé pour afficher une mire d'étalonnage, laquelle mire est réfléchi vers la webcam au moyen d'un miroir plan. Ensuite, en changeant le motif affiché à l'écran, nous avons montré que le dispositif pouvait être utilisé pour effectuer la reconstruction 3D par stéréophotométrie.

Ce travail préliminaire sera complété ultérieurement par une analyse plus profonde du modèle photométrique de l'image. Des évaluations quantitatives des deux étapes (étalonnage et reconstruction 3D) doivent également être effectuées.

Références

- [Bou04] BOUGUET J.-Y. : Camera calibration toolbox for matlab.
- [FY07] FUNK N., YANG Y.-H. : Using a raster display for photometric stereo. In *Computer and Robot Vision, 2007. CRV '07. Fourth Canadian Conference on* (2007), pp. 201–207.
- [HZ03] HARTLEY R., ZISSERMAN A. : *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [PSM*14] PARK J., SINHA S., MATSUSHITA Y., TAI Y., KWON I. : Calibrating a non-isotropic near point light source using a plane. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on* (2014), p. To appear.
- [RBN10] RODRIGUES R., BARRETO J., NUNES U. : Camera pose estimation using images of planar mirror reflections. In *European Conference on Computer Vision (CVPR 2010)* (2010).
- [Sch08] SCHINDLER G. : Photometric stereo via computer screen lighting for real-time surface reconstruction. In *International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)* (2008).
- [Woo80] WOODHAM R. J. : Photometric method for determining surface orientation from multiple images. *Optical Engineering*. Vol. 19, Num. 1 (1980), 139–144.